

Nonlinear system identification using integrated linear-NN models: series vs. parallel structures

H Zabiri, M. Ramasamy⁺, T D Lemma, A Maulud

Chemical Engineering Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750, Tronoh,
Perak, MALAYSIA

Abstract. In this paper, the performance of integrated linear-NN models is investigated for nonlinear system identification using two different structures: series vs. parallel. In particular, Laguerre filters are selected as the linear models, and multi-layer perceptron (MLP) or feed-forward neural networks (NN) are selected for the nonlinear models. Results show promising capability of the (novel) parallel Laguerre-NN structure especially in terms of its generalization capability when subjected to data different from those used during the identification stage in comparison to the series Laguerre-NN.

Keywords: nonlinear system identification, parallel integration, OBF, MLP, extrapolation

1. Introduction

In majority of control applications, system identification is an indispensable part for the analysis and controller design [1]. This applies for nonlinear systems as well, which has been the primary focus in recent years. One of the most frequently studied classes of nonlinear models consist of block-oriented representations, where linear dynamic systems and nonlinear static mappings, are represented by separate blocks connected in series [2].

Wiener model structure is one of the most commonly used block-oriented (BO) models and various applications in industrial nonlinear systems, such as distillation columns and pH processes have been reported [3-5]. Wiener models consist of a dynamic linear part cascaded with a static nonlinear component as shown in Figure 1. Common model classes for the dynamic linear subsystem are FIR filters, input/output models, state space models, and Orthonormal Basis Filters (OBF), e.g. Laguerre or Kautz filters. The memoryless requirement for the nonlinear subsystem allows almost unlimited choices for the nonlinear element, and recently the use of Neural Networks (NN) as the static nonlinear component has been investigated by various authors [3, 5]. In [3] especially, the separate blocks are represented by Laguerre filters and multi-layer perceptron (MLP) NN in series, respectively, and the models are solved sequentially. The Laguerre parameter 'p' is chosen based on the step response of the process at its nominal operating point, as 'p' dictates the dominant pole(s) of the process. After fixing a proper value of 'p' the network is then trained.

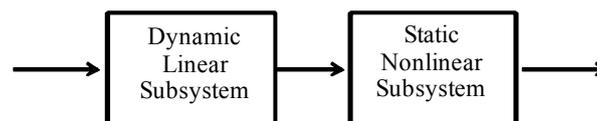


Figure 1. The block diagram of the Wiener model.

In contrast to the serial structure as discussed in [3], the parallel integration of linear-NN models approach as suggested in [6] provides an interesting alternative in modeling a nonlinear system (see Figure 2). This

⁺ Corresponding author. Tel.: +605-3687585; fax: +605-3656176.
E-mail address: marappagounder@petronas.com.my.

approach relies on the fact that a nonlinear model may perform worse than the linear one if it is not chosen appropriately. By developing a linear model (obtained by using either input/output data or by first principle strategy) in parallel with the nonlinear model, such that the overall model output is determined by the sum of the linear and the nonlinear parts, the performance of the overall nonlinear model is then ensured to be either equivalent or superior than the linear model. An example of this approach is presented in [7] for a vibration system, where a fundamental linear model is used in parallel with NN.

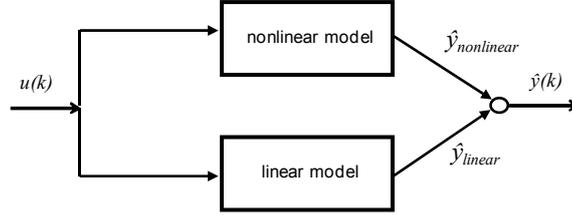


Figure 2. The schematic diagram of the parallel linear/nonlinear models [6].

In this paper, the performances of integrated Laguerre-NN models are investigated for both the series and parallel structures. In particular, the parallel structure to be presented in this paper is of novel type, which is different from that described in [7]. The series structure is developed based on [3]. The series and parallel Laguerre-NN models are investigated especially for its generalization capability in the presence of new data. In Section 2, the novel parallel structure is presented. The series structure is explained in detail in [3], for brevity purposes, interested readers are referred to the paper. To provide a fair comparison to the method in [3], MLP NN is selected. Results and discussions are presented in Section 3.

2. Novel parallel OBF-NN models

2.1. The idea

In prediction methods, the analysis of residuals is often underestimated [8]. There are occasions where residuals are not due to randomness and may actually inherit the characteristics of the original system. In this paper, the approach constitutes the usage of residuals in a sequential nonlinear identification technique using parallel integration of the linear OBF and nonlinear NN models.

The proposed configuration may be represented using a block diagram as shown in Figure 3. Note that in the proposed once-through sequential method, the linear dynamic OBF model is developed first. The input sequence to the NN is the original input sequence to the process, u , and the residuals (i.e. $e = y_m - y_1$, where y_m is the actual measured output and y_1 is the linear OBF model predicted output) are the output to be predicted by the nonlinear NN. Iterations, based on some convergence criteria of standard NN algorithm, are performed only in the NN block. The overall predicted output is then the summation of the predicted residuals, \hat{e} and the OBF output, y_1 .

Hence,

$$\hat{y}_i = y_1 + \hat{e} \quad (1)$$

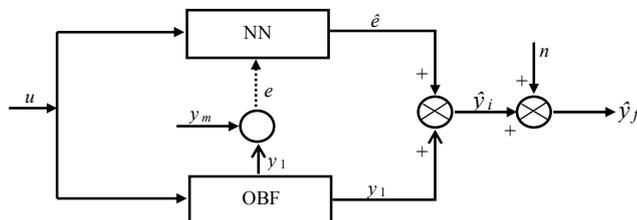


Figure 3. The block diagram of the proposed parallel OBF-NN models.

The measurement noise is considered to be acting on the output and is designated by n . The NN used throughout this paper is a single hidden layer MLP network. The mathematical descriptions of the methods are given below using Laguerre filters to represent the OBF models for ease of discussions.

2.2. Identification of parallel Laguerre-NN models

First, a parsimonious linear Laguerre model is identified using methods described in [9]. Given a set of nonlinear data to be identified $[u(k), y_m(k)]$, a Laguerre model is developed using crude estimate or arbitrarily chosen poles. Then, one or two of the dominant poles of the system are estimated using the methods

proposed in [9]. The estimated dominant poles are used to develop more accurate Laguerre model. A better estimate of the dominant poles is obtained from the new Laguerre model. The process is repeated until a convergence criterion is satisfied.

A SISO linear system modeled by Laguerre filters can generally be represented as follows [3]:

$$y_1(k) = \left(\sum_{i=1}^N c_i L_i(k) \right) u(k) \quad (2)$$

where

$$L_i(k) = \sqrt{(1-p^2)T_s} \frac{(1-pk)^{i-1}}{(k-p)^i} \quad (3)$$

$L_i(k)$ denotes the i^{th} order Laguerre filter, N the number of Laguerre filters used for model development, p the Laguerre filter parameters (i.e. the dominant poles of the system), T_s the sampling interval, $y_1(k)$ the model output, and $u(k)$ is the process input. Once a satisfactory Laguerre model is obtained, the model output y_1 is used to calculate the residuals. These residuals are then passed to the MLP NN. Considering a SISO example with a single output neuron, the output of an MLP neural network with one hidden layer is given by

$$\hat{e}(k) = \gamma(u(k), e(k)) = \beta \left[b^2 + \sum_{i=1}^K w_i^2 \varphi(b_i^1 + w_{i,1}^1 u(k)) \right] \quad (4)$$

where $\gamma(\cdot)$ is the nonlinear neural network function approximation, which is trained with data sets consisting of the input or regression vector, $u(k)$ and the residuals, $e(k)$. Also $\varphi, \beta: R \rightarrow R$ are the nonlinear activation functions (e.g. hyperbolic tangent etc.), b are the biases, K is the number of hidden neurons, and the weights of the network are denoted by $w_{i,j}^1, i=1, \dots, K$ (with i^{th} neuron and j^{th} input, in this case $j=1$) for the first layer, and $w_i^2, i=1, \dots, K$ for the second layer.

Without loss of generality and for ease of discussions, zero noise is assumed, i.e. $n=0$, such that $\hat{y}_f = \hat{y}_i$. The overall model output then is the summation of both the linear dynamic model and the predicted residuals \hat{e} .

$$\hat{y}_f(k) = y_1(k) + \hat{e}(k) \quad (5)$$

$$\hat{y}_f(k) = y_1(k) + \gamma(u(k), e(k)) \quad (6)$$

Substituting equations (2) and (4) yields,

$$\hat{y}_f(k) = \left(\sum_{j=1}^N c_j L_j(k) \right) u(k) + \beta \left[b^2 + \sum_{i=1}^K w_i^2 \varphi(b_i^1 + w_{i,1}^1 u(k)) \right] \quad (7)$$

2.3. Identification algorithm

The proposed identification algorithm is based on the recently developed identification algorithm for parsimonious OBF models by Lemma *et al.* [9] and the standard back-propagation (BP) algorithm for the neural network training. Given a set of nonlinear data to be identified $[u(k), y_m(k)]$, the algorithm can be described as follows:

1. Develop a parsimonious OBF model using methods described by [9] to get y_1 .
2. Calculate the residuals using $e = y_m - y_1$.
3. Develop the NN model using standard BP algorithm.

The input data to the NN is segregated as training and validation sets (75% for training, and 25% for validation) as is normally done with any NN modeling. The optimal NN model is determined by varying the activation functions for each layer and the regressors sets. Convergence criteria used is based on Root Mean Square Error (RMSE).

3. Results and Discussions

To demonstrate the performance of the series and parallel Laguerre-NN models, Van de Vusse reactor case study is considered [10] under medium nonlinearity condition. The nonlinear identification is carried out for SISO system by considering the dynamic characteristics from the changes of the feed flow rate, F , and the product outlet concentration, C_B . In all cases, the number of Laguerre filters is fixed at six (wherever applicable), and a single-hidden layer standard MLP network is adopted.

Figures 3 and 4 show the corresponding performance of the series and parallel methods when subjected to extrapolated data up to 30% increase in the feed flow rate, F , beyond the original training range used to identify the model. For comparison purposes, the performance of pure NN is also presented.

In the identification stage (Training set), all models, except for linear OBF, performs comparably similar as can be seen from the resulting RMSE values. However, when subjected to data that slowly drifting away from the original training set, it can be clearly observed that the novel parallel Laguerre-NN models have superior extrapolation performance in comparison to the series Laguerre-NN models as proposed in [3], as well as against pure NN. Pure NN behavior is as expected due to its constant extrapolation behavior. Series Laguerre-NN structure on the other hand, relies on pure linear behavior when subjected to extrapolation.

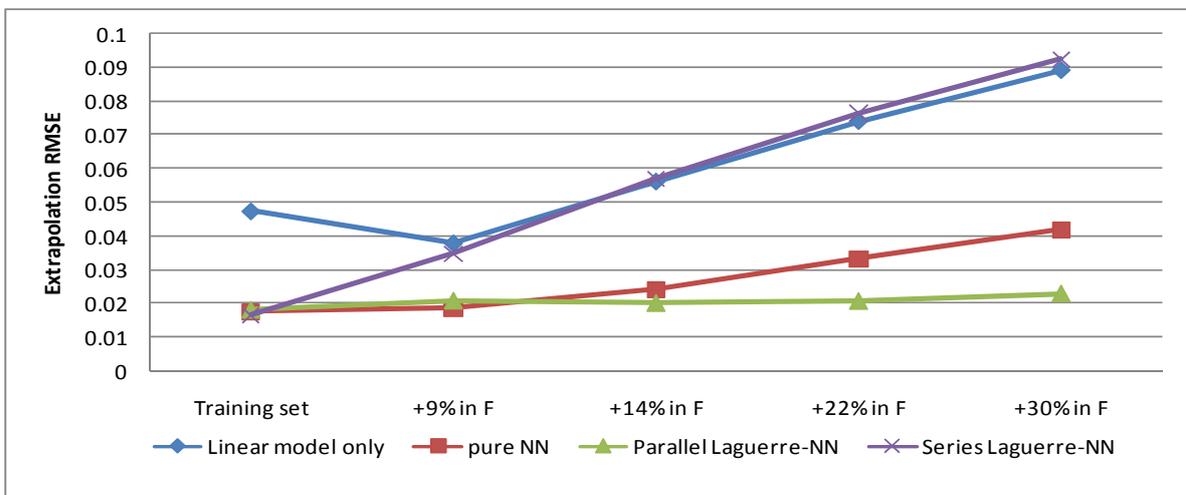
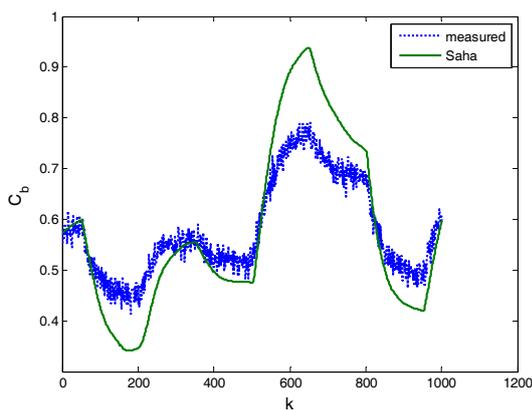
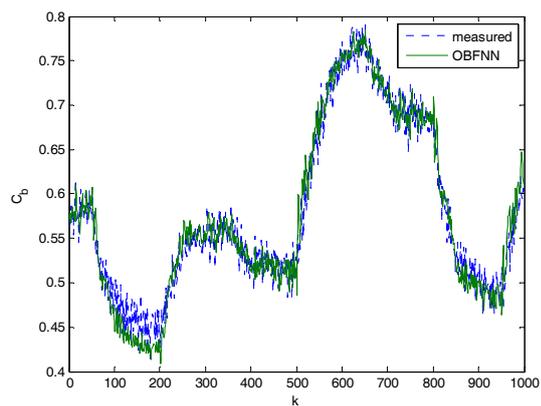


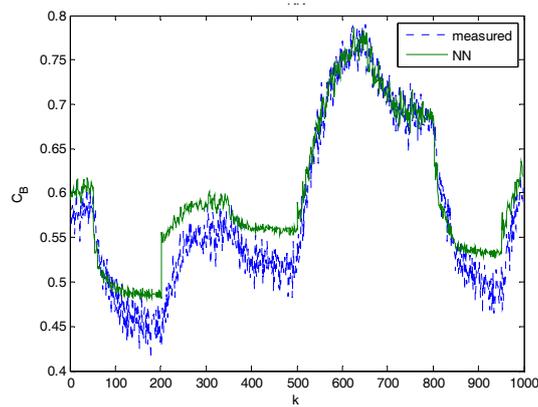
Figure 3. Root mean squared error (RMSE) for the training set and extrapolated data sets.



(a) Series Laguerre-NN [3]



(b) Parallel Laguerre-NN



(c) Pure NN

Figure 3. Performance comparison for a 22% increased in the inlet flow rate, F , beyond the original training range.

4. Conclusions

In this paper, the performance of series and parallel Laguerre-NN models are investigated. Using the nonlinear Van de Vusse reactor case study, it is shown that the novel parallel Laguerre-NN structure provides promising performance, with better extrapolation or generalization capability than the series Laguerre-NN model.

5. Acknowledgements

The authors would like to acknowledge Universiti Teknologi PETRONAS for the funding provided for this work.

6. References

- [1] Y.-L. Hsu, J.-S. Wang, "A Wiener-type recurrent neural network and its control strategy for nonlinear dynamic applications", *Journal of Process Control* 19 (2009) 942–953.
- [2] S. Totterman, H. T. Toivonen, "Support vector method for identification of Wiener models", *Journal of Process Control* (2009), 19, 1174-1181.
- [3] Saha, P., Krishnan, S.H., Rao, V.S.R., Patwardhan, S.C., 2004. "Modeling and predictive control of MIMO nonlinear systems using Wiener–Laguerre models". *Chemical Engineering Communication* 191, 1083-1119.
- [4] P. Aadaleesan, N. Miglan, R. Sharma, P. Saha, "Nonlinear system identification using Wiener-type Laguerre-Wavelet network model", *Chemical Engineering Science* (2008), 63, 3932-3941.
- [5] M. Lawrynczuk, "Computationally efficient nonlinear model predictive control based on neural Wiener models", *Neurocomputing* 74 (2010), 401-417.
- [6] Nelles, O., 2001. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer, Berlin.
- [7] T. Yasui, A. Moran, M. Hayase, "Integration of linear systems and neural networks for identification and control of nonlinear systems", *Proceedings of the 35th SICE Annual Conference (SICE'96)*, July, 24-26 1996, Tottori, 1389-1394.
- [8] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P. Glorennec, H. Hjalmarsen, A. Juditsky, "Nonlinear black-box modeling in system identification: a unified overview", *Automatica*, 31 (1995), 1691-1724.
- [9] Lemma, D. Tufa, M. Ramasamy, M. Shuhaimi, "Improved method for development of parsimonious orthonormal basis filter models", *Journal of Process Control*, 21 (2011), 36-45.
- [10] J. Vojtesek, P. Dostal, R. Haber, "Simulation and control of a continuous stirred tank reactor", In: *Proc. Of Sixth Portuguese Conference on Automatic Control CONTROLLO 2004*, Faro Portugal, 315-320.